

Dialect Variation in Stop Consonant Voicing

A Senior Honors Thesis

Presented in Partial Fulfillment of the Requirements for graduation
with research distinction in Speech and Hearing Sciences in the undergraduate colleges of
The Ohio State University

By

Samantha A. Lyle

The Ohio State University

June 2008

Project Advisors: Dr. Robert Fox and Dr. Ewa Jacewicz

ABSTRACT

Recent sociophonetic research has shown significant differences in the pronunciation of vowels among dialects. However, dialectal differences in stop consonant productions have not been as widely researched. This study examines the differences between speakers from south-central/southeastern Wisconsin and westernmost North Carolina, specifically in terms of the way the voiced stop /b/ is produced. Twenty female speakers were selected from recordings, ten from Wisconsin and ten from North Carolina. Each subject read two sets of thirty sentences that included five sets of target words. Acoustic measurements of the consonant /b/ and the target word itself were completed. From these measurements the following variables were calculated: closure duration, word duration, proportion of closure duration to word duration, duration of voicing during closure, proportion of voicing in closure and frequency of complete voicing through closure. The results of these analyses show there are significant differences in the ways the consonant /b/ is produced in Wisconsin and North Carolina. The greatest differences were found in the total duration of voicing during consonant closure, proportion of voicing in closure and the proportion of times the stop closure was completely voiced. The present results provide a comprehensive set of data for a detailed dialect comparison of stop production in these two dialects of American English.

TABLE OF CONTENTS

Chapter I. Introduction and Literature Review.....	3
Chapter II. Methodology.....	13
Chapter III. Results.....	18
Chapter IV. Discussion.....	29
Chapter V. Acknowledgments.....	33
Chapter VI. References.....	34
Index of Figures.....	37
Appendix.....	38

CHAPTER 1

INTRODUCTION AND LITERATURE REVIEW

Human interaction through vocal communication is the result of a phenomenon known as speech. It is through speech that many people express thoughts, feelings, and information with one another. The speech mechanism is complex and consists of many different systems that must work together in order to produce vocal speech. A very important factor in speaking is the production of voice itself. Voicing, or phonation, is often overlooked and taken for granted in this elaborate system.

1.1 Production of Voicing

In the larynx there is a pair of vocal folds that attach anteriorly inside the thyroid cartilage and posteriorly to the left and right arytenoid cartilages. These vocal folds separate two cavities in the vocal tract, the mouth and lungs. The space between these two folds is called the glottis. When the vocal folds are abducted, they are apart and the glottis is open. When the vocal folds are adducted, they are together and the airway to the lungs is sealed. Vocal folds must be adducted in order for voicing to occur.

In speech there are voiced and voiceless sounds. Voiceless sounds have an absence of vocal fold vibration, or glottal buzz. Voiced sounds are characterized by the presence of vocal fold vibration, producing periodicity in the speech wave. Below the vocal folds is the subglottal airway, and above the vocal folds is the supraglottal airway. In order for voicing to occur there must be a pressure difference between these airways because there has to be airflow in order for phonation to occur. Subglottal pressure

increases during exhalation because of the decrease in volume when the thoracic cavity is compressed. When the subglottal pressure is greater than the supraglottal pressure the result is a pressure drop across the glottis. According to the myoelastic-aerodynamic theory, the vocal folds will only vibrate when this pressure drop across the glottis is present. Since fluids travel from high pressure to low pressure the air pushes up against the vocal folds forcing them apart.

The lower portion of the vocal folds separate first while the upper portion remains together. As the air continues to travel, the upper portion of the vocal folds begin to separate, opening the glottis. The elasticity of the vocal folds first brings the lower portion of the vocal folds back to a medial position, which lowers the transglottal pressure. This decrease in pressure pulls the lower portion of the vocal folds inward, with the upper portion lagging behind. As the upper portion of the vocal folds comes together the glottis is sealed. The vocal folds rapidly cycle through this vibratory pattern producing phonation. Phonation is known as the generation of sound due to the vibration of vocal folds. (Behrman, 2007)

1.2 Acoustic Characteristics of Stop Consonants

Oral stop consonants are produced by a complete occlusion in the oral cavity where airflow is briefly yet completely stopped. When oral stops are produced the velopharyngeal port is closed preventing air from escaping through the nasal passages. These stops are also called plosives because of the burst of sound that exists after the constriction of air is released. The articulatory production is not the only thing that can define oral stop consonants; they can also be described by their acoustic characteristics.

There are both voiceless and voiced stops in American English. The voiceless stops are /p,t,k/ and the voiced stops are /b,d,g/. While producing the plosive, preceding the release, there is a period of time called stop closure (representing that portion of the production when the oral cavity is closed) that may either be silent for voiceless stops or have low amplitude voicing for voiced stops.

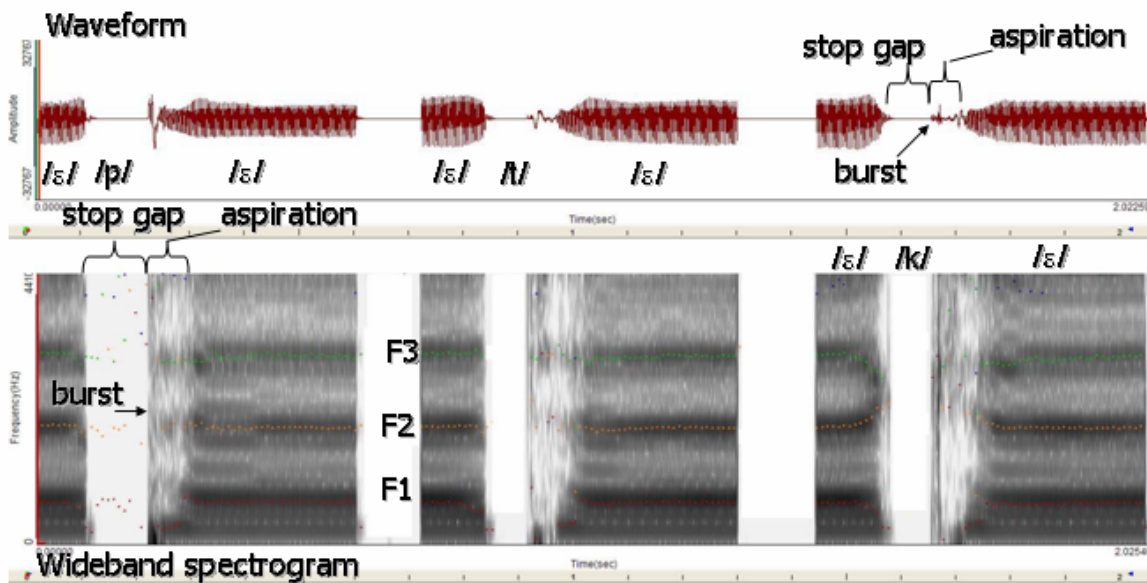


Figure 1.1 Waveform (top) and a wideband spectrogram (bottom) of the voiceless stops /p, t, k/ produced in an intervocalic position. (Behrman, 2007)

Figure 1.1 shows the waveform and a wideband spectrogram of three voiceless stop consonants, /p,t,k/, which occur in an intervocalic position. In the voiceless stops there is a period of complete closure where no voicing is present, or the “stop gap”. When the increased intraoral pressure meets the atmospheric pressure upon release of the constriction, the result is a sudden burst. The acoustic consequence of this burst, or closure release, is a spike in amplitude of a relatively broadband noise transient that is

evident in the waveform and the spectrogram. There is a period of aspiration that follows the closure release. This is present among many voiceless stops in English, especially when the stop is in the word initial or intervocalic position. Aspiration is defined as a brief hiss of air, or a breathy noise, and there is usually low amplitude or no voicing present until the onset of the vowel. Aspiration is produced as air flows through vocal folds, which are partially closed, into the pharynx. This noise sounds much like the glottal fricative /h/, like in the word *hot*. The aspiration, or breath of air immediately follows the closure release. For example, when you put your hand up to your mouth and say the word “push” you will feel the puff of air following the /p/. In word-final position the stop is usually unreleased and no aspiration occurs.

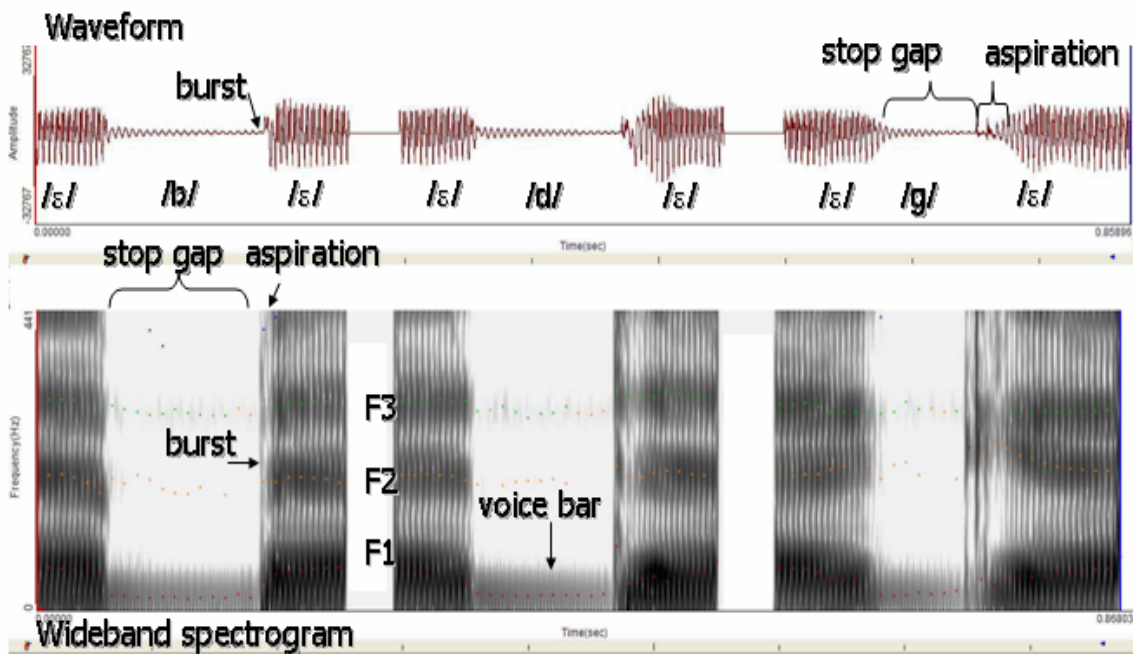


Figure 1.2 Waveform (top) and a wideband spectrogram (bottom) of the voiced stops /b, d, g/ produced in an intervocalic position. (Behrman, 2007)

Figure 1.2 shows the waveform and a wideband spectrogram of the voiced stops /b,d,g/. In voiced stops there is not a silent closure like in voiceless stops. Voiced stops usually have low amplitude voicing, due to damping, with varying amounts of silence. In the waveform, the voicing is seen as a quasi-periodic signal which is also present in the spectrogram in the form of a voice bar. Aspiration is rarely seen in voiced stops.

One property of stop consonants has been particularly well studied, that of voice onset time (VOT). VOT is defined as the time interval between the articulatory release of a stop and the onset of vocal fold vibration which signals a beginning of a vowel that follows a stop consonant (Kent and Read, 1992). VOT can be recognized by a number of acoustic cues such as quasi-periodic energy following the burst, amplitude changes during the noise burst, and the existence of F1 cutback (Lisker & Abramson, 1964). VOT has been found to be an effective means to distinguish between voicing categories in oral stops. For example, the value of VOT is a good indicator of voiced and voiceless stops in English. Voiceless stops have the so called long voicing lag, ranging from 20 ms to 80 ms. Voiced stops in English have a short voicing lag which can range from 2ms to 20 ms. This difference can be seen in Figure 1.3, in which the time interval for VOT in voiceless stops is much longer than in voiced stops.

VOT may also be negative and simultaneous. Negative VOT indicates that voicing occurs before the release of the stop consonant, which is also called prevoicing. Simultaneous VOT is when the release of the stop consonant occurs at the same time as the onset of the periodic voicing of the vowel. Some languages use these properties to signal phonemic distinctions. Studies have shown that speaking rate has an influence on VOT and affects VOT in voiceless stops more than in voiced stops. For example,

Kessinger and Blumstein (1998) found that as speaking rate slows in voiceless stops, VOT and vowel duration equally lengthen. Research suggests that VOTs tend to be longer when they precede a high vowel rather than preceding a low vowel. This may, in part, be due to the reduced transglottal pressure drop that can cause vocal fold vibration to cease which is a result of the constriction of the sonorant and high vowel (Behrman, 2007).

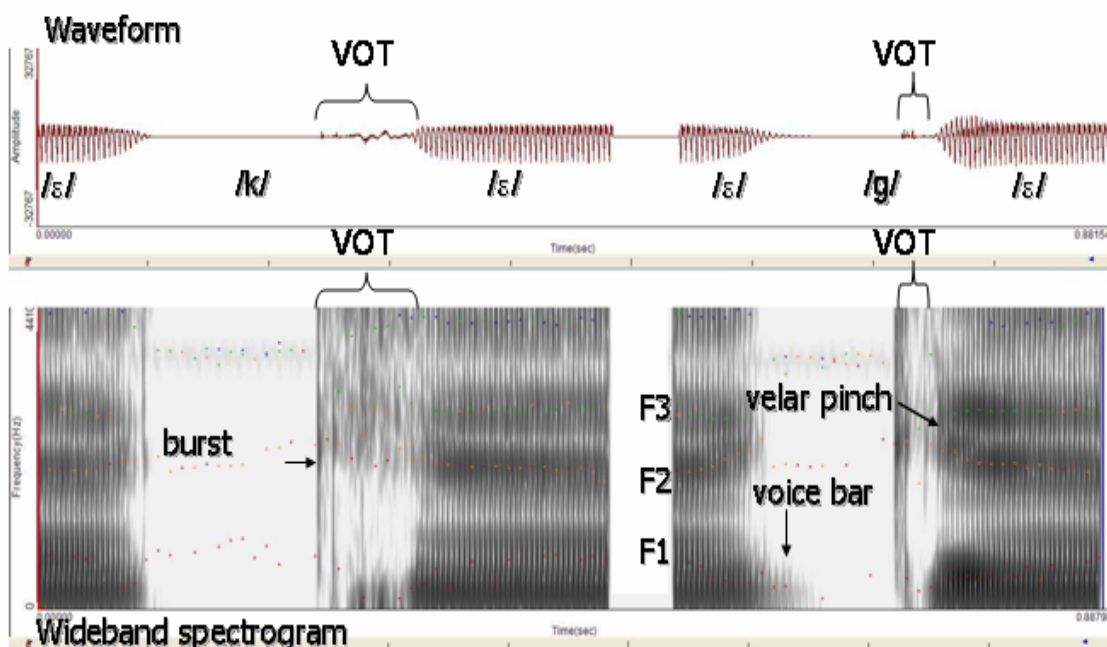


Figure 1.3 Waveform (top) and a wideband spectrogram (bottom) of the voiceless stop /k/ and the voiced stop /g/ showing the differences in voice onset time. (Behrman, 2007)

1.3 Variation in stop closure voicing

It is known that there are variations of acoustic properties within speech among different languages. Variations in stop closure voicing can be affected by several factors such as position of the stop consonant within the sequence of speech segments (e.g., phonetic context or syllable structure) or non-phonetic factors such as speaking rate and

speaker characteristics. One source of variability, phonetic context, has been found especially effective in changing the degree of voicing of stop consonants. The primary positions of stops are prevocalic, intervocalic, and postvocalic. Prevocalic position is when the consonant immediately precedes a vowel. Intervocalic means the consonant occurs between two vowels, both immediately following one and preceding another vowel in a speech stream. Postvocalic position is when the consonant follows the vowel.

Westbury and Keating (1986) examined whether it is ‘natural’ for stop consonants to be voiced or voiceless in different phonetic contexts. They found that if speakers actually do seek out the easiest or most ‘natural’ way to produce sound sequences, then they would minimize change in articulatory parameters. Doing this suggests that a stop in the medial, or intervocalic, position should largely be voiced as long as the closure duration is short. If the closure is long, intervocalic stops tend to show a voiced-voiceless pattern. Westbury and Keating state it is more likely for stop voicing to occur in the medial/intervocalic position than initial or final position due largely to the fact that voicing depends on difference between subglottal and supraglottal pressures. There appears to be a greater difference in these pressures when the stop consonant occurs in the medial position, rather than the initial or final position, which makes voicing more ‘natural’ in this position.

This finding is further explored in the current study. As it will become apparent, all stop consonants examined in this study are either between voiced sonorants or in an intervocalic position. These positions, along with the variation in emphasis of the target word, which may also affect stop closure voicing, provide an excellent testing ground for measuring the degree of voicing in voiced stops.

1.4 Language vs. Dialect

There are evident variations in voicing across different languages. Keating *et al* (1983) surveyed several languages for their study and found that eighteen of fifty-one displayed some sort of neutralization in regards to voicing-related contrasts among stops. Some languages favor voiceless unaspirated stops in the medial position although speakers must exert a greater articulatory effort to do so. Research shows that at a phonemic level, voiceless stops are largely preferred over voiced stops across languages.

Cho and Ladefoged (1999) studied the VOT of speakers from 18 different languages. They recognize that VOT may vary with place of articulation. They state that there is a longer VOT when the closure is further back and there is a more extended contact area. They also state that VOT is shorter with faster movements of articulators. They found that some differences in VOT between languages can be explained by physiological and aerodynamic causes where others require language specific explanations. Cho and Ladefoged state that regardless of articulatory gestures, languages still have unpredictable variations. They recognize three “universally specified” values of VOT which are voiced, voiceless unaspirated and aspirated.

Observing that there are many known variations across languages in regards to voicing characteristics, there is a legitimate question whether these variations exist across dialects of the same language. It is well established that there are dialectal differences concerning vowel quality as well as place and manner of articulation of consonants. For example in African American Vernacular English the phoneme /f/ is often substituted for /θ/ and in the Bostonian dialect the postvocalic /r/ does not exist. In some Southern dialects /z/ and /ð/ are often neutralized and become /d/, for example *wasn't* becomes

wadn't and *them* becomes *dem* (Wolfram & Schilling-Estes, 2006). Dialect studies primarily focus on variations in vowel systems and examine vowel changes and shifts (see Labov, 1994, for a review). Consonant productions, in regards to dialects in North America, have been researched but not as instrumentally as vowels. So the question that arises is “Are there significant differences in voicing characteristics in stop consonants among different American English dialects?”

1.5 Purpose of Study

This thesis examines the possible phonetic differences in the way voiced stops may be produced by speakers of two very different regional varieties of American English: westernmost North Carolina (Appalachian English) and south-central/southeastern Wisconsin (Inland North). These two dialects differ greatly in the phonetic characteristics of their vowels and sociophonetic research has shown significant differences in the pronunciation of vowels among these dialects (Jacewicz et al., 2006; 2007). In general, vowels have been shown to be the primary factor producing the distinct regional “accents” (e.g., southern or northern accents). However, as already mentioned, there is not as much research on dialectal differences in consonant production. This study is an effort to fill in this gap in research. The large differences in acoustic characteristics of “northern” and “southern” vowels let us expect at least some differences in the acoustics of “northern” and “southern” stop consonants. Of specific interest is the variation in the way the voiced stop /b/ is pronounced in these two varieties of American English. Some of the possibilities that exist include a complete voicing all the way through the stop closure, a partial voicing of the closure, and differences in the

length of VOT. The study will also examine possible temporal differences such as closure duration, word duration in which the stop consonant occurs, and proportion of voicing during the stop closure.

CHAPTER 2

METHODOLOGY

2.1 Speakers

Twenty adult female speakers were recorded for the experiment. Ten speakers were from south-central Wisconsin (the Madison area: Dodge and Dane counties) and ten were from western North Carolina (the Sylva, Cullowhee, and Waynesville areas: Jackson, Swain and Haywood counties). All speakers were born, raised, and spent most of their lives in the respective areas. They ranged in age from 51 to 65 years, and were paid volunteers. Each speaker was paid \$15.00 for her participation in a recording session which lasted approximately an hour to an hour and fifteen minutes.

2.2 Stimuli

The stimuli included target words that were measured from samples of controlled speech. The structure of these target words were /bVts/ and /bVdz/, where V represents one of the following target vowels: /ɪ, ɛ, æ, e, aɪ/. Speakers read two sets of thirty sentences that included five sets of target words such as *bits/bids*, *bets/beds*, *bats/bads*, *baits/bades*, and *bites/bides*. These target words were produced in the same sentential and phonetic context (between voiced sonorants). This means that the target consonant in each word is produced between voiced sounds. In this study the target consonant /b/ is produced following the voiced sonorant /l/ and before a vowel. However, prosodic variations were systematically introduced for each set to create different emphasis conditions. The three levels of emphasis of the target word are high, intermediate and low. This variation in emphasis was obtained by varying the main sentence stress. The

word that was to be emphasized in each sentence was capitalized on the screen for the reader to see. Examples of these sentences include:

Bits

HIGH

John knows the small SCREWS are sharp.

No! John knows the small **BITS** are sharp.

INTERMEDIATE

John knows the SOFT bits are sharp.

No! John knows the SMALL **bits** are sharp.

LOW

John knows the small bits are DULL.

No! John knows the small **bits** are SHARP

Bids

HIGH

Ted thinks the fall SALES are low.

No! Ted thinks the fall **BIDS** are low.

INTERMEDIATE

Ted thinks the SPRING bids are low.

No! Ted thinks the FALL **bids** are low.

LOW

Ted thinks the fall bids are HIGH.

No! Ted thinks the fall **bids** are LOW.

For this study, only the target word from the second sentence in the set was examined

(seen here in bold print). These are only examples of the vowel /I/ in the words “bits”

and “bids”. For a full list of these sentences see Appendix A.

2.3 Recording Procedure

Recording of sentences was controlled by a custom program written in Matlab. The sentence pairs were randomized and appeared on a computer monitor. The participant read each sentence pair speaking into a head-mounted microphone (Shure SM10A), placed approximately 1-inch distance from the lips. The sentences were recorded directly onto a hard drive disc at a sampling rate of 44.1 kHz. The experimenter only accepted fluently read sentences with proper emphasis placement. The recordings were repeated as many times as needed to obtain adequate productions. Two research assistants helped with collection of data, one in Wisconsin and one in North Carolina, and all participants in a given state were recorded by the same experimenter.

2.4 Acoustic Measurements

Acoustic measurements of the consonant /b/ and the target word itself were completed to identify and mark a set of acoustic landmarks including the stop closure onset, closure release, voicing offset during the stop closure, voicing onset for the vowel, word onset (which was the same as the stop closure onset) and word offset. Measurements were made by hand from the waveform (with reference to the spectrogram) using Adobe Audition 1.0 speech analysis program. A Matlab program was then used to display all the waveforms with the markings that were made to check to make sure they were correct. A second check of all acoustic landmarks was performed by a research advisor.

Stop closure onset was located at the zero-crossing (crossing of the x-axis) where acoustic energy of the preceding sonorant consonant was significantly decreased and

when there was a change in periodicity which signaled a beginning of a stop closure. The closure release was located at the zero-crossing where there was a burst of acoustic energy for the release of the stop closure. The voicing offset during the closure (if present) was located where acoustic energy and periodicity ceased. Vowel onset was located after the closure release at the zero-crossing of the first vertical striation, or glottal pulse of voicing. The location of word onset was the same as the location of stop closure onset, (i.e. the measurement of stop closure onset was taken as the beginning of the word). Word offset was located at the end of the frication noise of the fricative that followed the second stop as in “bids” or “bits”.

From these measurements the following variables were calculated: word duration, closure duration, proportion of closure duration to word duration, duration of voicing during closure, proportion of voicing in closure, VOT and frequency of occurrence of complete voicing through closure in the whole data set. Closure duration was calculated (in milliseconds) by subtracting the stop closure onset from the closure release. The word duration was calculated (in milliseconds) by subtracting the word onset from the word offset. The proportion of closure duration to word duration was a ratio of these two measures. The duration of voicing during closure was calculated (in milliseconds) by subtracting the stop closure onset from the voicing offset if it existed. If the closure was completely voiced the duration of voicing equaled the closure duration. The proportion of voicing in the closure is a percentage of how much of the closure duration has voicing, for example if it was voiced throughout it would be 100%. Frequency of voicing through closure is a proportion of the amount of times the closure was voiced throughout to total

number of closures (reported here as a percentage). VOT was calculated by subtracting the closure release from the onset of voicing for the vowel.

2.5 Statistical Analysis

Repeated measures analyses of variance (ANOVAs) were conducted on word duration, stop closure duration, proportion of closure-to-word duration, closure voicing duration, percentage of closure voicing, frequency of a voiced-through closure, and VOT. The within-subject factors were final consonant in the word (/t/ or /d/), emphasis position (high, intermediate, low) and vowel (/ɪ, ɛ, æ, e, aɪ/). Dialect was the between-subject factor. In addition to the significance values, a measure of the effect size – partial eta squared (η^2) – is also reported. The value of η^2 can range from 0.0 to 1.0 and it should be considered a measure of the proportion of variance explained by a dependent variable when controlling for other factors.

CHAPTER 3

RESULTS

3.1 RESULTS

Before presenting the results for each measure examined in this study, it may be useful to provide a few examples showing the nature of variation in closure voicing for a typical Wisconsin and a typical North Carolina speaker analyzed here.

Figure 3.1 shows waveforms of closures of the stop /b/ in the words *bad*es and *ba*its produced by a 55-year old female Wisconsin speaker and a 59-year old female North Carolina speaker. Both speakers read the set of sentences for this study with comparable fluency (i.e. there were no pauses in their productions) and at a comparable articulation rate, which was 3.23 syll/s for the Wisconsin speaker and 3.17 syll/s for the North Carolina speaker. The waveform displays include stop closures for each emphasis level examined here. The displays are time aligned so that each waveform begins with a 15-ms final portion of /l/ preceding the stop closure. The closure terminates with a second 15-ms interval measured from the release, which consists of release burst (if present) and a portion of vowel onset.

As can be seen, there is a clear difference between the closures of the WI and NC speaker. All NC closures are fully voiced whereas WI closures begin with a period of voicing which ceases gradually and the closure terminates in a complete silence. There is a clearly marked release burst for this particular WI speaker whereas no such release can be detected in the production of the NC speaker. The longest closure was found in the high emphasis position of the word, followed by intermediate and low positions,

respectively, although the difference between the latter two positions is rather small. The WI closures tend to be longer than NC closures across all emphasis levels.

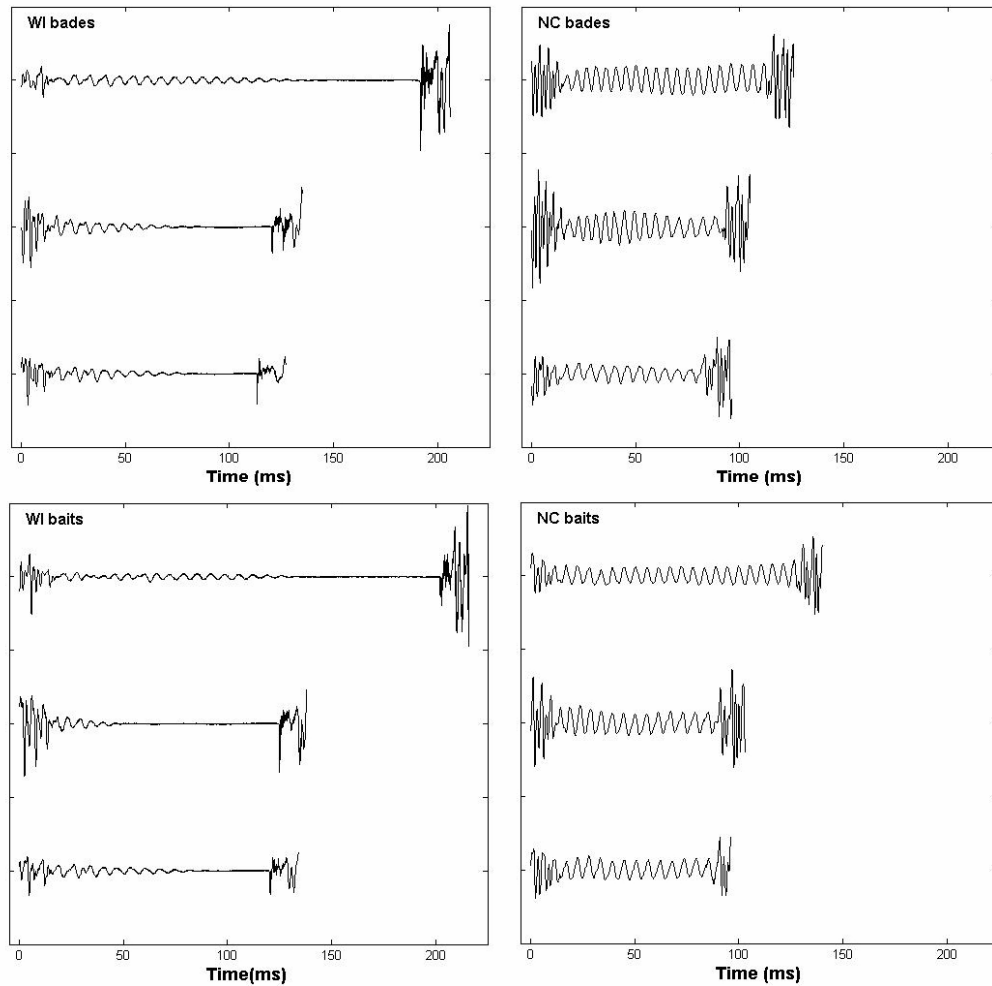


Figure 3.1 The left side panels are waveforms of a Wisconsin production of the words *bades* and *bait*. The right side panels are waveforms of a North Carolina production of the words *bades* and *bait*.

With these differences in mind, the results are now presented for each measure selected in this study to assess the general trend and significance of differences between the two types of stop closures. First, the variation in word duration and stop closure duration are examined to determine a proportion of closure duration to word duration, which may vary cross-dialectally and may contribute to the nature of closure voicing itself. Next, the closure voicing is explored by assessing its duration during the closure,

its proportion in closure, and the frequency with which the voiced-through closure occurred in the present sample. Finally, VOT (if present) is examined for the two types of closures.

3.2 Word duration

The overall mean word duration was 422 ms for WI speakers and 464 ms for NC speakers. On average, NC words were 9% longer. However, the ANOVA results showed no significant main effect of dialect, which indicates that this difference needs to be regarded as a tendency rather than a true dialectal effect. All three within-subject factors were significant. The main effect of final consonant indicated that words in the *b_d* context were significantly longer than in the *b_t* context ($F(1, 18) = 10.25, p = 0.005, \eta^2 = 0.363$). The strong significant effect of emphasis position ($F(1.6, 28) = 101.26, p < 0.001, \eta^2 = 0.849$), showed that, on average, words in high emphasis position were longest (534 ms), followed by intermediate (427 ms) and low positions (367 ms), respectively. As shown in Figure 3.2, these differences were well represented across all WI and NC instances of the target words.

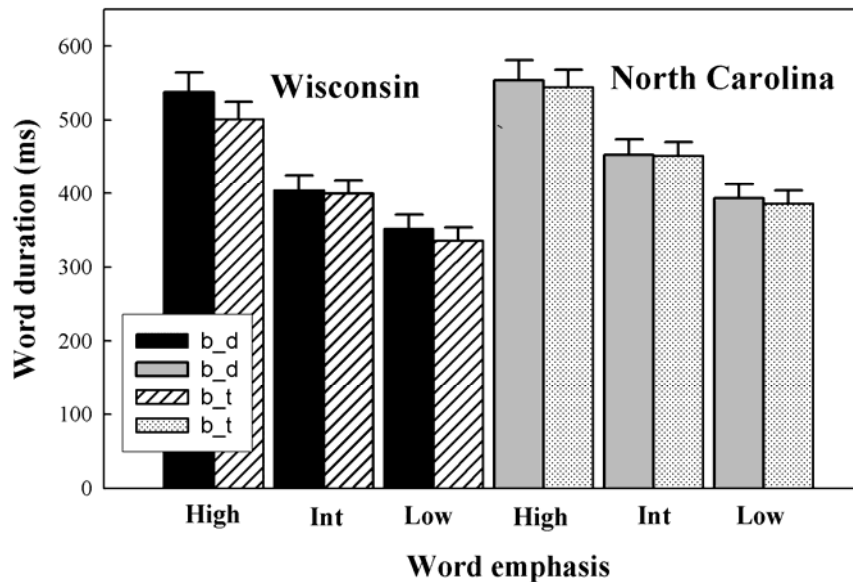


Figure 3.2 Relationship between word duration, final consonant, emphasis position and dialect.

There was also a strong significant main effect of vowel ($F(3.5, 63.3) = 50.33, p < 0.001, \eta^2 = 0.737$). Words containing one of the short vowels /ɪ, ɛ/ were on average shorter (412 and 420 ms, respectively) than words containing longer or diphthongal vowels /e, æ, aɪ/ (457, 461 and 463 ms, respectively).

3.3 Stop closure duration

The analysis of closure duration intended to assess its variability, which was expected given that the words were produced with different levels of emphasis and contained different vowel categories. On average, stop closure was longer for WI speakers than for NC speakers (110 ms vs. 101 ms). The main effect of dialect was not significant, however. There was a significant main effect of emphasis position ($F(1.9, 33.3) = 40.26, p < 0.001, \eta^2 = 0.691$) indicating that mean closure duration was longest when the word was highly emphasized and gradually decreased in the intermediate and low emphasis positions. The mean duration values, in descending order, were 137, 99 and 80 ms. The main effect of vowel was also significant ($F(3, 54.1) = 6.18, p = 0.001, \eta^2 = 0.256$). There was no clear relation between the length of the closure and duration of a particular vowel category. The longest closure (mean 110 ms) was found for the vowel /ɛ/ and the shortest was for the vowel /æ/ (mean 98 ms). Significant was also the three way interaction between final consonant, stressed position and vowel ($F(3.6, 64.6) = 3.55, p = 0.014, \eta^2 = 0.165$), which is illustrated for each dialect separately in Figure 3.3.

As Figure 3.3 shows, the relation between closure duration, vowel category, emphasis position, and final consonant in the word is very complex. The closure duration varies greatly as a function of all these factors although it is noteworthy that the degree of emphasis affects the stop closure duration in a systematic way across all vowel categories and word types examined here.

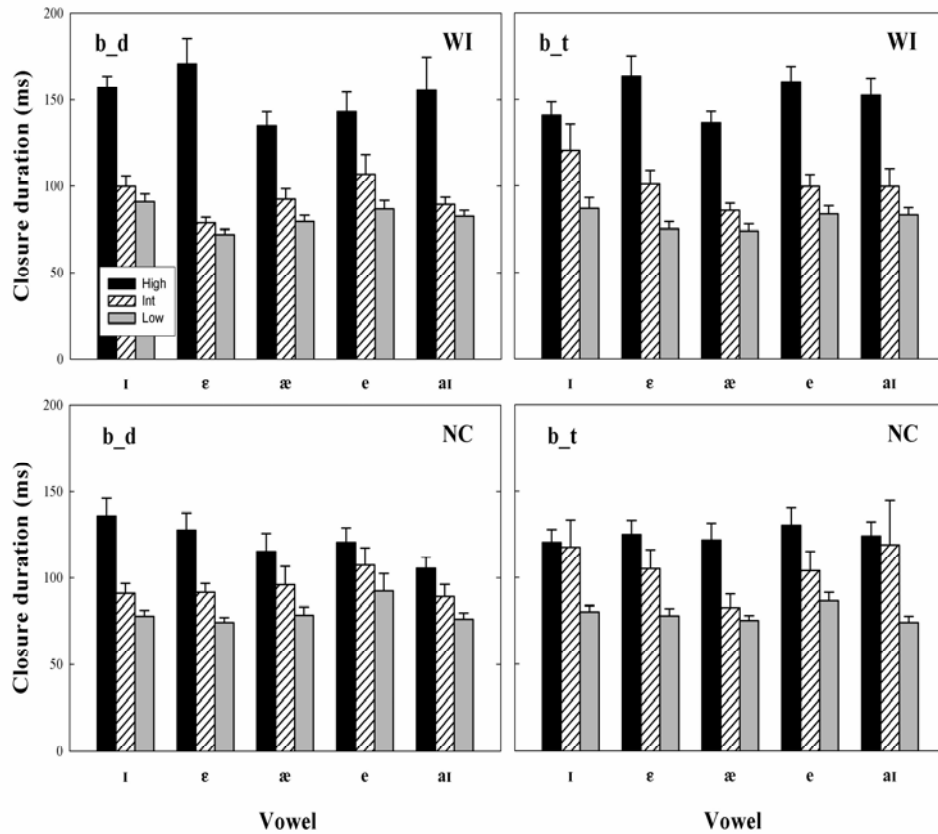


Figure 3.3 Closure duration split by both dialect and final consonant. Within each panel the varying degrees of emphasis are shown.

3.4 Proportion of closure-to-word duration

Although closure duration, measured in absolute terms, was generally longer for WI speakers, the effects of dialect were more pronounced when closure duration was expressed in relative terms, i.e. as a proportion to the duration of the word (with values ranging from 0 to 100%). For WI speakers, the mean proportion of closure-to-word

duration was greater (26%) than for NC speakers (22%) and the main effect of dialect was significant ($F(1, 18) = 6.5, p = 0.020, \eta^2 = 0.264$). The proportion of closure duration was significantly greater in *b_t* words than in *b_d* words ($F(1, 18) = 34.7, p < 0.001, \eta^2 = 0.658$) and varied significantly as a function of emphasis position ($F(1.4, 25.5) = 7.3, p = 0.006, \eta^2 = 0.289$). The mean proportion of closure duration was greatest in the high emphasis position (26%) followed by intermediate and low, respectively (23 vs. 22%). The strong significant effect of vowel ($F(3.8, 69.2) = 59.03, p < 0.001, \eta^2 = 0.766$) indicated that the proportion of closure duration was greatest when the words contained short vowels /ɪ, ɛ/ (27 and 25%, respectively), followed by diphthongal vowels /e, aɪ/ (24 and 22%, respectively) and the vowel /æ/ (21%). A significant vowel by dialect interaction ($F(3.8, 69.2) = 3.74, p = 0.009, \eta^2 = 0.172$) indicated, however, that this seemingly straightforward relation varies as a function of dialect, which is illustrated in Figure 3.4.

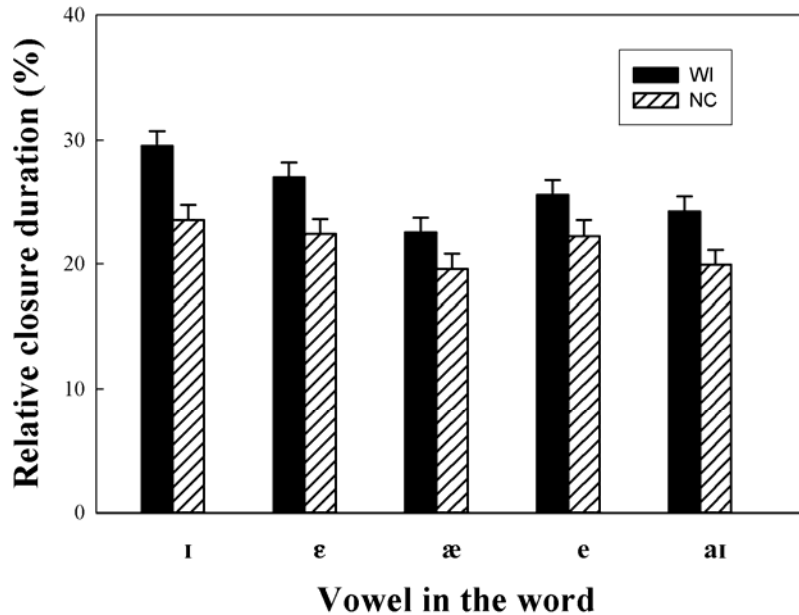


Figure 3.4. Proportion of closure duration for each vowel and for each dialect.

As can be seen, the proportions of closure duration for words containing NC variants of /æ/ and /aɪ/ are equally low. Similarly, there is no difference in the proportion of closure duration for words containing the NC vowels /ɛ/ and /e/. These differences between the two dialects let us expect dialect-specific variations in the duration of the voicing period during the consonant closure. The question arises whether the voicing portion of the closure is also longer for Wisconsin speakers, whose proportion of closure duration to word duration is greater than for North Carolina speakers.

3.5 Closure voicing duration

As it turned out, mean closure voicing duration for WI speakers was shorter than for NC speakers (69 vs. 89 ms) and the main effect of dialect was significant ($F(1, 18) = 6.43, p = 0.021, \eta^2 = 0.263$). Significant was also the effect of final consonant in the word ($F(1, 18) = 9.34, p = 0.007, \eta^2 = 0.342$), showing that the voicing portion of the closure was shorter in *b_t* words as compared to *b_d* words. Closure voicing duration also varied significantly as a function of word emphasis ($F(1.8, 31.8) = 27, p < 0.001, \eta^2 = 0.600$), indicating that voicing was more extensive in high emphasis position, followed by intermediate and low, respectively. Figure 3.5 illustrates the effects of word emphasis on closure voicing duration for each dialect.

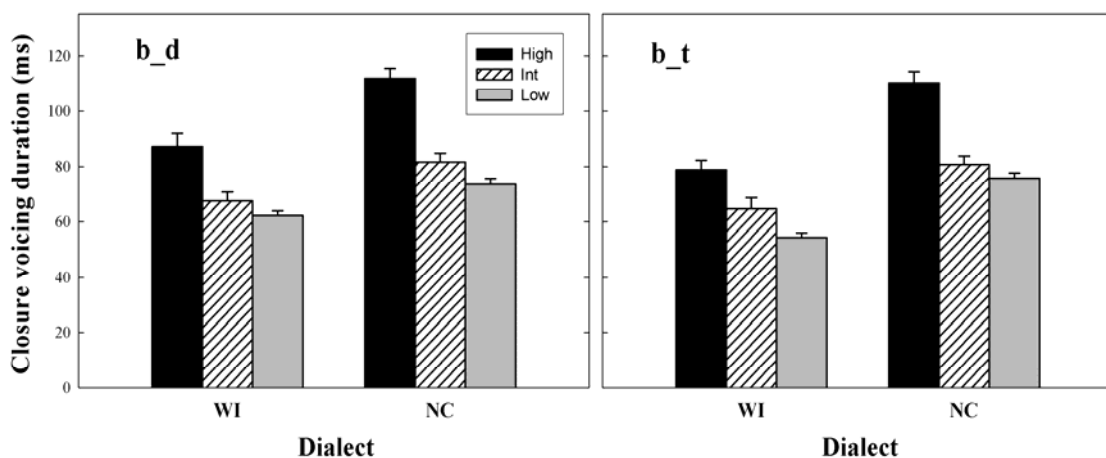


Figure 3.5 Closure voicing duration for each dialect as a function of vowel emphasis.

Of particular interest is a significant interaction between the final consonant and dialect ($F(1, 18) = 8.65, p = 0.009, \eta^2 = 0.324$). This interaction shows that voicing duration is shorter in *b_t* than in *b_d* words for WI speakers (66 vs. 72 ms) but not for NC speakers (89 and 89 ms). We can interpret this result as a kind of anticipatory effect for WI speakers whose expectation of a voiceless stop in word final position is manifested in their less extensive voicing of the word initial /b/. This effect was not found for NC speakers.

3.6 Proportion of voicing in closure

Given the differences in closure duration for WI and NC speakers, a cross-dialectal comparison of closure voicing is more direct when the voicing portion is assessed relative to the duration of the closure. An analysis of the proportion of voicing in closure showed a significant effect of dialect ($F(1, 18) = 18.1, p < 0.001, \eta^2 = 0.501$). Proportion of voicing in closure was on average smaller for WI speakers than for NC speakers (67 vs. 92%), indicating that NC variant of /b/ is almost entirely voiced during the stop closure. The main effects of both final consonant and emphasis position were significant ($F(1, 18) = 34.1, p < 0.001, \eta^2 = 0.655$ and $F(1.5, 26.5) = 9.6, p = 0.002, \eta^2 = 0.349$, respectively). However, it was the significant interactions between final consonant and dialect and between emphasis position and dialect that shed more light on the dialect-specific changes in the proportion of closure voicing as a function of either within-subject factor.

In particular, the final consonant by dialect interaction ($F(1, 18) = 14.2, p = 0.001, \eta^2 = 0.441$) showed that the proportion of closure voicing was greater in b_d words than in b_t words for WI speakers (72 vs. 65%) but not for NC speakers (92 vs. 91%). The second interaction, that of emphasis position by dialect ($F(1.5, 26.5) = 5.8, p = 0.014, \eta^2 = 0.243$) is illustrated in Figure 5. As can be seen, there are clear dialectal differences in the proportion of closure voicing. For WI speakers, there is a relationship between the length of the closure and the proportion of voicing: The shorter the closure duration, the greater the proportion of voicing in closure. Thus, proportion of closure voicing is smallest in words in high emphasis position, followed by intermediate and low, respectively. For NC speakers, there is no such relationship and the closure is almost entirely voiced regardless of the variation in closure duration.

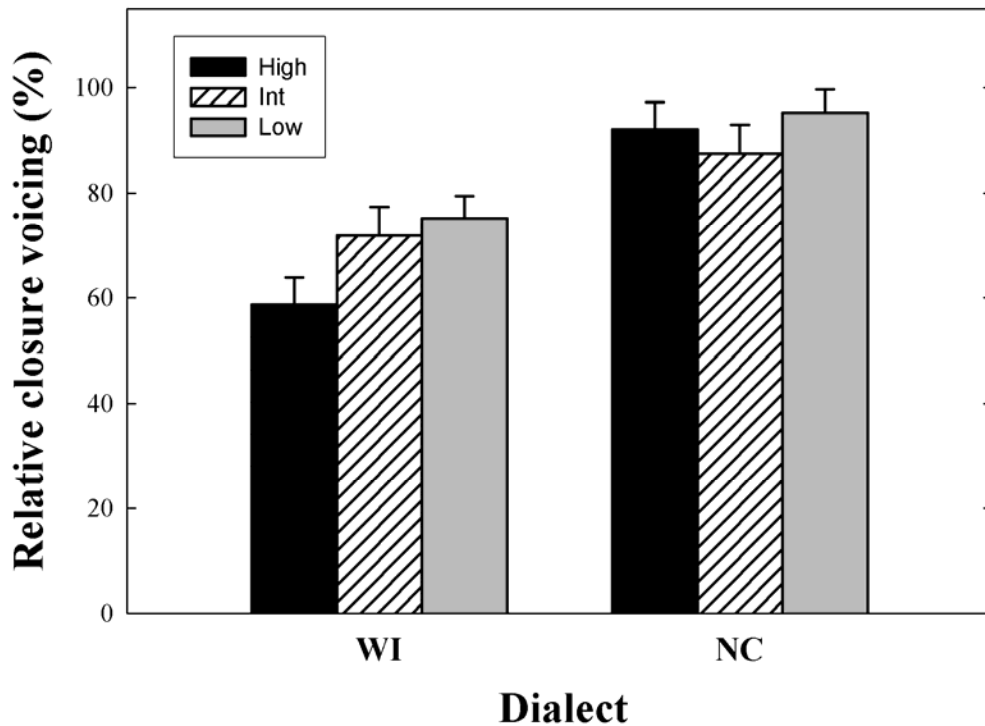


Figure 3.6 Proportion of voicing in closure for each dialect as a function of vowel emphasis.

3.7 Proportion of the voiced-through closure

We also examined the proportion of the voiced-through closure in the present data set for NC and WI stops. The results show a strong effect of dialect ($F(1, 18) = 24.8, p < 0.001, \eta^2 = 0.580$), indicating that the proportion of voiced-through closures was significantly greater for NC speakers (73%) than for WI speakers (24%). There was also a significant main effect of final consonant ($F(1, 18) = 8.9, p = 0.008, \eta^2 = 0.331$), showing that a voiced-through closure occurred more often in *b_d* words than in *b_t* words (52 vs. 45%). Finally, there was significant effect of emphasis position ($F(1.9, 33.9) = 6.84, p = 0.004, \eta^2 = 0.275$). The proportion of voiced-through closures was greatest in the low emphasis position (59%) and decreased in intermediate and high positions, respectively (46% and 41%).

This measure clearly shows that a voiced-through closure occurs more often when there is a condition to reduce closure duration such as presence of a voiceless final consonant in the target word or decrease in word emphasis.

3.8 Voice onset time (VOT)

Given the significant dialectal differences in the nature of closure voicing, we also expected dialectal differences in the VOT. In this analysis, we excluded all cases of fully voiced closures because there is no “voicing onset” event that can be identified. We included in this analysis only those tokens in which voicing was stopped at some point during the closure. Two separate independent samples *t*-tests were applied to *b_t* and *b_d* words. For *b_t* words, the effect of dialect was significant ($t = 3.58, df = 64.79, p = 0.001$). Mean VOT value (in ms) for WI speakers was 4.13 whereas for NC speakers it

was -5.39, indicating prevoicing of /b/ in NC but not in WI variants. A similar result was found for *b_d* words ($t = 3.65$, $df = 50.42$, $p = 0.001$) although the prevoicing for NC speakers was even longer (WI mean was 3.95 ms and NC mean was -13.31 ms). Given the significant disparity between the two dialects in the overall numbers of stop closures that were fully voiced (NC speakers had many more fully voiced stops than WI speakers); one cannot assume equal variances in comparisons of means using *t*-tests. Therefore, all relevant *t*-tests were completed assuming non-equal variances which increased the estimate of the standard error and produced a more conservative test.

CHAPTER 4

DISCUSSION

Before this study, little research has been conducted to analyze the differences between dialects and their consonant productions. In previous research it is evident that there are phonetic differences across different languages, and even in sociophonetic research there are significant differences in vowels across dialects. However, consonant production variations may prove to be a strong source of insight for differences between dialects. A systematic variation in the production of the stop was introduced by varying the degree of emphasis of the target word beginning with the stop, vowel quality and the status of voicing of the word-final consonant cluster. Considering all these factors, the results provided an explanation for the impressionistic perception stated at the outset that North Carolina speakers seem to produce more sonorous variants of the stops as compared to Wisconsin speakers.

The current study was successful in analyzing how the stop consonant /b/ is produced in two different regional dialects. There are significant differences between these two dialects. The Wisconsin dialect is consistent with the suggestions of Westbury and Keating in that intervocalic stops are naturally voiced as long as the closure duration is short. Wisconsin had more fully voiced closures in the low emphasis position, which was the shortest closure duration. In the high emphasis position, Wisconsin speakers showed a voiced-voiceless pattern, which is also consistent with Westbury and Keating. However, North Carolina speakers generally deviate from this statement because their voicing patterns were not affected by change in emphasis position or closure duration.

In general, we found Wisconsin speakers producing stops with longer closures despite shorter word durations as compared to North Carolina speakers. The closure duration differences were greatest when the target word was in the high emphasis position and tended to diminish with each position of lower emphasis. Wisconsin speakers showed that the proportion of closure to word duration was greater for words that ended in voiceless consonants rather than voiced, North Carolina speakers did not show this pattern. The effects of vowel category were not consistent and no clear pattern was detected.

The Wisconsin closures were usually not fully voiced and the average voicing portion of the closure did not last longer than 67% of closure duration. These closures, that were not fully voiced, terminated in silence and were followed by a closure release. North Carolina speakers had a higher proportion of closure voicing which reached an average of 92%. Emphasis position also had significant effects on the proportion of closure voicing. For Wisconsin speakers, words that were produced with high emphasis had the smallest proportion of closure voicing whereas low emphasis positions had the greatest proportion. For North Carolina speakers there was no such pattern.

Perhaps the most dramatic dialectal difference is the number of closures that were fully voiced. North Carolina speakers produced the majority of the fully voiced closures in the present sample. For Wisconsin speakers, the fully voiced closures were sparse and occurred mostly in the low emphasis positions.

Clearly, these two different patterns of stop closure voicing for NC and WI speakers come from differences in the way voicing is maintained during the closure by the speech production mechanism. As stated before, voicing is produced by vocal fold

vibration, which can only occur if there is adequate transglottal pressure. We hypothesize that the transglottal pressure for Wisconsin speakers terminates early, which decreases and then terminates the amplitude in voicing. The North Carolina speakers demonstrated a different way of maintaining voicing during the stop closure. The closures were mostly fully voiced and the proportion of voicing during the closure was generally not sensitive to the variation in closure duration as a function of word emphasis. It appears that North Carolina speakers were able to maintain transglottal pressure during the stop closure by additional articulatory maneuvers, most likely by lowering the velum and venting the air through the nose.

Because this study involves acoustic analysis only and no aerodynamic data are available for the present set of acoustic measurements we cannot assume with certainty that North Carolina speakers utilize the velum to sustain the voicing during the stop closure. However, the sound quality of the stop itself and of the speech from the majority of our North Carolina speakers in general gives us an indication that the velopharyngeal port is at least open partially allowing air to escape through the nasal tract. Appalachian speech has long been described (and stereotyped) as having at least some degree of nasality present even in words that have no nasal segments. This is often called a "nasal twang". Further studies involving aerodynamics would need to be performed to confirm this statement.

The current findings should be explored further with future research. There are limited resources regarding the effects of dialect on stop consonant voicing. Other variables of interest would be the effects of age and gender on the dialects. It would

also be instructive to determine whether these dialectal patterns are maintained in unconstrained informal speech.

CHAPTER 5

ACKNOWLEDGEMENTS

This project was supported by The Ohio State University College of Arts and Sciences and the College of Social and Behavioral Sciences.

I would like to thank Dr. Robert Allen Fox and Dr. Ewa Jacewicz for their patience and assistance with this thesis.

I would like to acknowledge Dr. Brian D. Joseph for serving on my defense committee as well as everyone who worked in SPA Labs for their support.

CHAPTER 6

REFERENCES

Behrman, Alison. *Speech and Voice Science*. 1st ed. Plural Inc., 2007.

Cho, Taihong, & Ladefoged, P. (1999). Variations and universals in VOT: evidence from 18 languages. *Journal of Phonetics* 27. 207-229.

Jacewicz, E., Fox, R. A., & Salmons, J. (2006). Prosodic prominence effects on vowels in chain shifts. *Language Variation And Change*. 18 (3), 285-316.

Jacewicz, E., Salmons, J., & Fox, R. (2007). Vowel Duration in Three American English Dialects. *American Speech*. 82 (4), 367-385.

Keating, P., Linker, W., & Huffman, M. (1983). Closure duration of stop consonants. *Journal of Phonetics*, 11, 277-290.

Kent, R. D., & Read, C. (1992). *The acoustic analysis of speech*. San Diego, Calif: Singular Pub. Group.

Kessinger, R., & Blumstein, S. (1997). Effects of speaking rate on voice-onset time in Thai, French, and English. *Journal of Phonetics*, 25, 143-168.

Kessinger, R., & Blumstein, S. (1998). Effects of speaking rate on voice-onset time and vowel production: Some implications for perception studies. *Journal of Phonetics*, 26, 117-128.

Koenig, L., & Lucero, J. (2008). Stop consonant voicing and intraoral pressure contours in women and children. *Journal of the Acoustical Society of America*, 123 (2), 1077-1088.

Labov, T. (1994). H. Varenne, Ambiguous harmony: Family talk in America. *Language in Society*. 23 (1), 124.

Lisker, L. & Abramson, A.S. (1964). A cross-language study of voicing in initial stops: acoustical measurements. *Word*, 20, 384-422.

Pickett, J.M. The acoustics of speech communication: fundamentals, speech perception theory, and technology. Needham Heights: Allyn & Bacon, 1999.

Pind, J. (1999). The role of F1 in the perception of voice onset time and voice offset time. *Journal of the Acoustical Society of America* 106, 434-437.

Purnell, T., Salmons, J., Tepeli, D., & Mercer, J. (2005). Structured Heterogeneity and Change in Laryngeal Phonetics: Upper Midwestern Final Obstruents. *Journal of English Linguistics* 33 (4), 307-338.

Ryalls, J., Zipprer, A., & Baldauff, P. (1997). A preliminary investigation of the effects of gender and race on voice onset time. *Journal of Speech, Language, and Hearing Research*, 40, 642-645.

Westbury, J., & Keating, P. (1985). On the naturalness of stop consonant voicing. *Journal of Linguistics* 22, 145-166.

Wolfram, Walt, and Natalie Schilling-Estes. American English Dialects and Variation. 2nd ed. Malden: Blackwell, 2006. 81.

Yoshioka, H., Lofqvist, A., & Hirose, H. (1981). Laryngeal adjustments on the production of consonant clusters and geminates in American English. *Journal of the Acoustical Society of America* 70 (6), 1615-1623.

Index of Figures

Figure 1.1.....	5
Figure 1.2.....	6
Figure 1.3.....	8
Figure 3.1.....	19
Figure 3.2.....	20
Figure 3.3.....	22
Figure 3.4.....	23
Figure 3.5.....	24
Figure 3.6.....	26

APPENDIX

The following sets of sentences were recorded by each speaker. All 2-set sentences were randomly presented to the subject in two stimulus lists.

Vowels before a voiceless consonant in a word

bits

John knows the SOFT bits are sharp.
No! John knows the SMALL bits are sharp.

John knows the small SCREWS are sharp.
No! John knows the small BITS are sharp.

John knows the small bits are DULL.
No! John knows the small bits are SHARP.

baits

Dad said the BRIGHT baits are best.
No! Dad said the DULL baits are best.

Dad said the dull HOOKS are best.
No! Dad said the dull BAITS are best.

Dad said the dull baits are WORST.
No! Dad said the dull baits are BEST.

bets

John said the BIG bets are low.
No! John said the SMALL bets are low.

John said the small POTS are low.
No! John said the small BETS are low.

John said the small bets are HIGH.
No! John said the small bets are LOW.

bats

Doc said the LARGE bats are fast.

No! Doc said the SMALL bats are fast.

Doc said the small BIRDS are fast.
No! Doc said the small BATS are fast.

Doc said the small bats are SLOW.
No! Doc said the small bats are FAST.

bites

Sue thinks the LARGE bites are deep.
No! Sue thinks the SMALL bites are deep.

Sue thinks the small CUTS are deep.
No! Sue thinks the small BITES are deep.

Sue thinks the small bites are WIDE.
No! Sue thinks the small bites are DEEP.

Vowels before a voiced consonant in a word

bids

Ted thinks the SPRING bids are low.
No! Ted thinks the FALL bids are low.

Ted thinks the fall SALES are low.
No! Ted thinks the fall BIDS are low.

Ted thinks the fall bids are HIGH.
No! Ted thinks the fall bids are LOW.

bades

(The nonsense word *bade* was explained to the speaker as indicating “a brand of knife, a brand name.”)

Ted says the SHARP bades are cheap.
No! Ted says the DULL bades are cheap.

Ted says the dull FORKS are cheap.
No! Ted says the dull BADES are cheap.

Ted says the dull bades are WEAK.
No! Ted says the dull bades are CHEAP.

beds

Rob said the SHORT beds are warm.
No! Rob said the TALL beds are warm.

Rob said the tall CHAIRS are warm.
No! Rob said the tall BEDS are warm.

Rob said the tall beds are COLD.
No! Rob said the tall beds are WARM.

bads

(The speaker was told that *bad* refers to “an error or mistake.” For example, if someone makes an error, he or she might say “my bad” instead of “my mistake.”).

Mike thinks the BIG bads are worse.
No! Mike thinks the SMALL bads are worse.

Mike thinks the small GOODS are worse.
No! Mike thinks the small BADS are worse.

Mike thinks the small bads are BEST.
No! Mike thinks the small bads are WORSE.

bids

(The nonsense word *bide* was explained to the speaker as indicating “a small animal, a type of dog.”)

Jane thinks the SHORT bides are cute.
No! Jane thinks the TALL bides are cute.

Jane thinks the small CATS are cute.
No! Jane thinks the small BIDES are cute.

Jane thinks the small bides are GROSS.
No! Jane thinks the small bides are CUTE.